MA677 Hypothesis Testing Assignment

Xihao Cao, Jack Carbaugh, Guangze Yu, Handing Zhang

2/14/2022

Hypothesis Testing

The goal of the hypothesis test is to evaluate a claim when given sample data from a population. The test involves checking if the sample data supports the claim, or if there is no evidence of it. The null hypothesis acts as the basis or previously held belief that we are looking to disprove, while the alternative hypothesis is the claim.

In the example for this report, it is previous knowledge that an ordinary aspirin is effective against headaches 60% of the time. A company's claim is that their new aspirin is more effective than the ordinary aspiring at relieving headaches. Thus, we have are testing the effectiveness of the new aspirin using sample data. If the proportion of headache relief, p, is not greater than the old aspirin, then we would expect $p = 0.6$. This will be the null hypothesis. The company's claim is that $p > 0.6$, so this will be the alternative hypothesis.

To make a decision about whether to reject the null hypothesis or not, we need a critical value of the proportion of relieved headaches to act as a decision boundary. The critical value must be selected to avoid error in the testing. If the critical value is too low, we are more likely to commit a type 1 error, which is the probability of the test being a false positive. For this problem, that would mean claiming the new aspirin has a sizeable effect when it actually does not, leading the company to spend money on and produce a faulty product. However, if the critical value is too high, we are more likely to commit a type 2 error, which is the probability of the test being a false negative. For this problem, that would mean not having enough evidence to claim the new aspirin has a sizeable effect when in reality it does, meaning the company missed an opportunity on an investment. Power is the probability of not making a type 2 error, so we would like to have a critical value that gives the test high power.

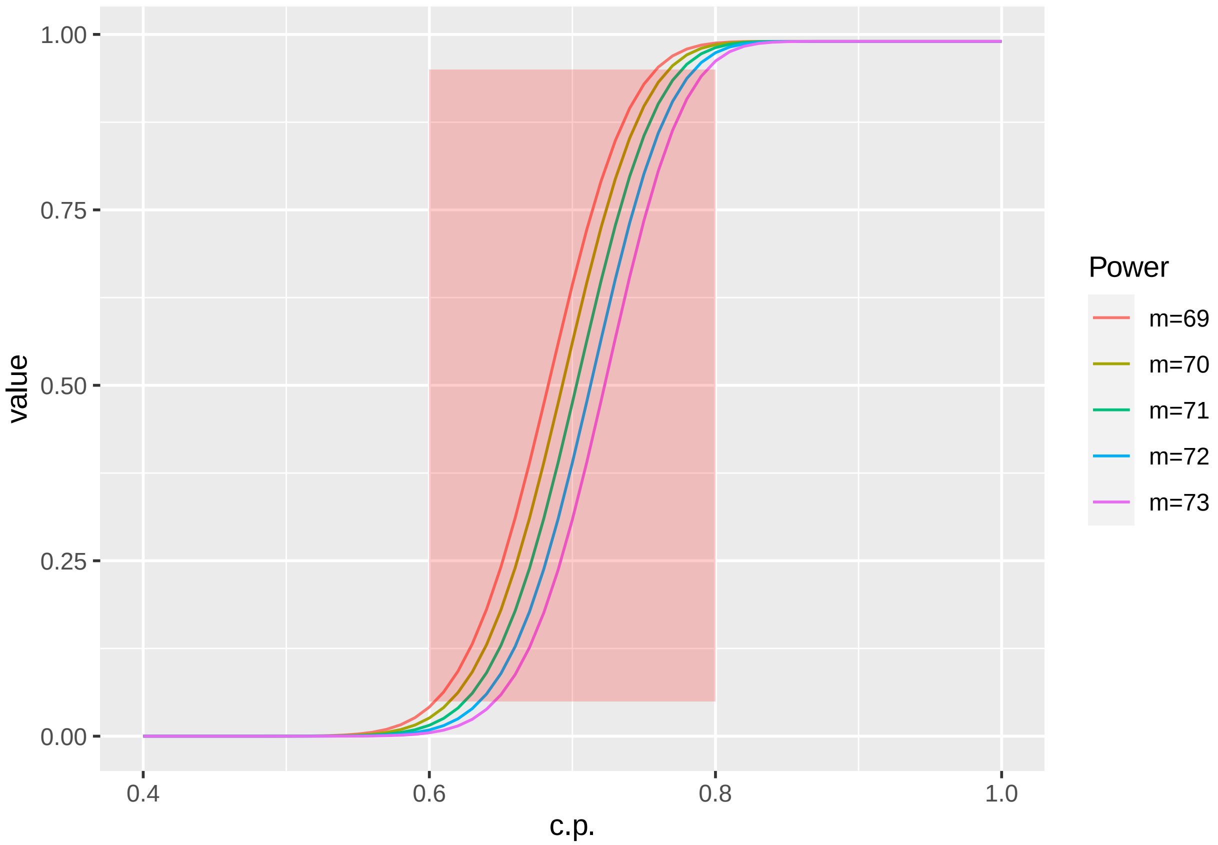
For this case, we want to have a critical value so that both type 1 and type 2 error is below 5% for the test. The analysis below attempts to locate a good critical value.

```r
m <- seq(69,73,1)
n <- 100
p <- seq(0.4, 1, 1/n)

result <- data.frame(c(p))
for (i in m) {
    Power <- cumsum(dbinom(i,n,p))
    result[,ncol(result)+1] <- Power
    names(result)[ncol(result)] <- paste("m=",i,sep="")
}

df <- melt(result,id.vars = 'c.p.', variable.name = 'Power')
```

```r
ggplot(df, aes(c.p.,value)) +
  geom_line(aes(colour = Power)) +
  annotate("rect", xmin=0.6, xmax=0.8, ymin=0.05, ymax=0.95, alpha=0.2, fill="red")
```



The plot shows the critical values that achieve type 1 and type 2 error below 5%. The x-axis is the true population proportion, the y-axis is the probability of type 1 error, and the different lines represent results at different critical values. The curve passing through the bottom of the box means that the critical value passes our type 1 criterion, while passing through the top means it has passed our type 2 criterion. m = 69 - 73 were the only values able to satisfy both requirements.